Experiments # 1

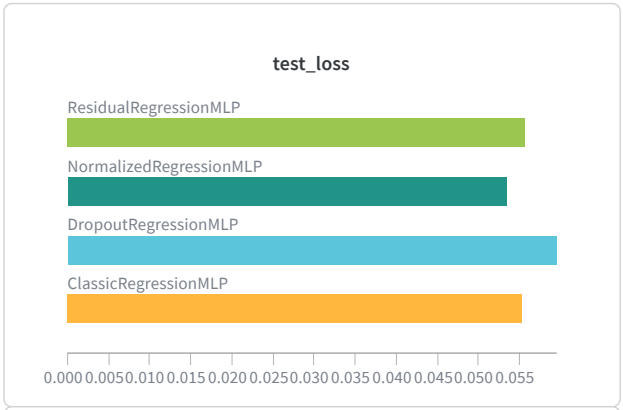
Experiments with architectures

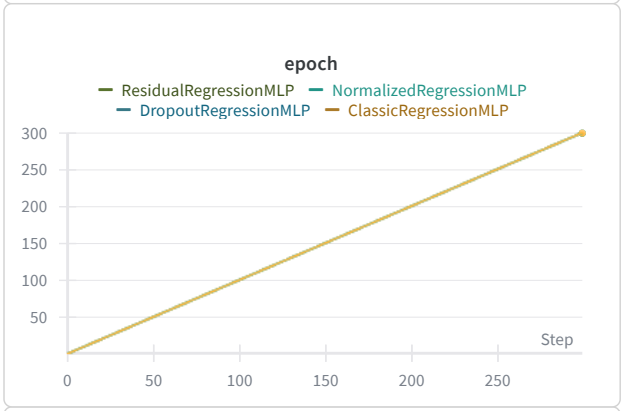
Nikolai Dobrydnev, Milan Seliga +

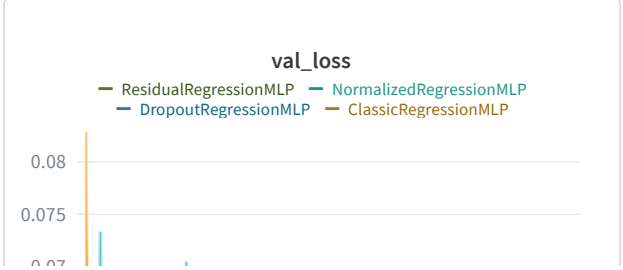
We tried to change the structure of our neural network by adding or changing some layers. The first model used standard fully connected layers and ReLu activation between them. In the second model, we tried adding dropout layers, but the resulting loss was greater than when using the first model. We then tried adding normalization to the first model, which resulted in the model learning faster and the smallest loss. Finally, we also tried residual architecture (with normalizations), but the resulting loss was slightly greater than when only normalization was added.

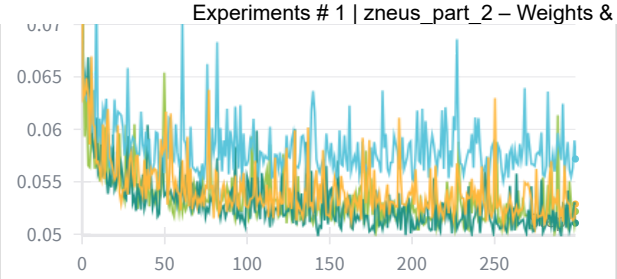
In the subsequent search for hyperparameters (using wandb sweep grid search), we therefore used a model with normalization.

Visual Support








0.065
0.06
0.055
0.05
0
50
100
150
200
250